An Interactive Learning and Adaptation Framework for Adaptive Robot Assisted Therapy

Konstantinos Tsiakas
HERACLEIA Lab
CSE Department
University of Texas at Arlington
konstantinos.tsiakas@mavs.uta.edu

Michalis Papakostas
HERACLEIA Lab
CSE Department
University of Texas at Arlington
michalis.papakostas@mavs.uta.edu

Benjamin Chebaa
HERACLEIA Lab
Department of Psychology
University of Texas at Arlington
benjamin.chebaa@mavs.uta.edu

Dylan Ebert
HERACLEIA Lab
CSE Department
University of Texas at Arlington
dylan.ebert@mavs.uta.edu

Vangelis Karkaletsis
National Center for Scientific
Research Demokritos,
Athens, Greece
vangelis@iit.demokritos.gr

Fillia Makedon
HERACLEIA Lab
CSE Department
University of Texas at Arlington
makedon@uta.edu

ABSTRACT

In this paper, we present an interactive learning and adaptation framework. The framework combines Interactive Reinforcement Learning methods to effectively adapt and refine a learned policy to cope with new users. We argue that implicit feedback provided by the primary user and guidance from a secondary user can be integrated to the adaptation mechanism, resulting at a tailored and safe interaction. We illustrate this framework with a use case in Robot Assisted Therapy, presenting a Robot Yoga Trainer that monitors a yoga training session, adjusting the session parameters based on human motion activity recognition and evaluation through depth data, to assist the user complete the session, following a Reinforcement Learning approach.

CCS Concepts

•Computing methodologies → Reinforcement learning; Learning from implicit feedback; Activity recognition and understanding;

Keywords

Interactive Reinforcement Learning, Policy Adaptation, Adaptive Robot Assisted Therapy

1. INTRODUCTION

Interactive Learning Agents are entities that learn what to do, through interacting with their environment – *agent policy*. A significant attribute of these agents is the adaptability of their policy towards a goal, in a dynamic and stochastic environment, as when a human user is involved in the interaction [8]. For this reason, interactive agents and systems have been successfully employed to Robot Assisted Therapy, with applications to physical and cognitive rehabilitation [24, 15].

Two major aspects of such applications are *safety* and *personalization*. An intelligent interactive agent should be adaptive to each user needs, preferences and abilities [16], while ensuring a safe interaction [5]. Such applications may also support the participation of a *secondary user*, who supervises the interaction with the *primary user*, resulting to a multiparty interaction (child, robot and therapist [6]). In real-world applications, such agents need to interact with many users. Thus, an agent should be able to adapt to different users by efficiently modifying a learned policy, instead of learning from scratch for each user [27].

In this work, we present an interactive learning framework that combines *Interactive Reinforcement Learning* [22, 1] and *Transfer Learning* [21, 27] methods to facilitate the policy adaptation of an agent to new users. We argue that interactive learning techniques can be used for adaptation, exploiting the expertise of an expert that supervises and guides the interaction, when needed, as well as the *implicit* feedback, provided by the primary user, in the form of an *affective signal* (e.g, heart rate).

We illustrate the proposed framework with an application in Robot Assisted Therapy, presenting a Robot Yoga Trainer that adjusts the session parameters to assist the user complete the session efficiently. The system evaluates the user's performance, through a human activity recognition and evaluation module. The main focus of this paper is to investigate how such a system can modify an existing policy towards a new user, exploiting additional communication channels as feedback and guidance.

PETRA '16, June 29-July 01, 2016, Corfu Island, Greece

DOI: http://dx.doi.org/10.1145/2910674.2935857

2. BACKGROUND AND RELATED WORK

Robot Assisted Therapy has been widely observed and tested as a tool to advance physical rehabilitation in specific cases, such as upper limb function for individuals that suffered from a stroke, or music therapy for people with dementia [20, 14]. Among this research, there is a consistent theme for the need to address patient motivation and engagement [13].

Reinforcement Learning (RL) provides an appropriate framework for interaction modeling and optimization and has been successfully applied to interactive systems with applications to HRI systems [15, 24]. In [20], the authors present a reinforcement learning approach for a long-term learning and adaptive socially assistive robotic system. Their main focus is policy adaptation for optimizing basic interaction parameters as proxemics and vocal content, so as to refine it to the user personality, preferences and needs and thus improve the user task performance.

In [17], the authors present a study on robot assisted tutoring to children, discussing the need of providing adaptive support to each child during a tutoring interaction. They argue that the adaptive robot should use affective feedback, provided by the child, to further tailor its support strategies, enabling the agent to interact with different users. In [18], they propose a supervised autonomy method that enables the therapist, as a secondary user, to guide the early interaction steps of the interaction is a robot assisted therapy application, where random exploration-based learning is not desirable, since it may lead to harmful interaction.

Interactive Reinforcement Learning (IRL) is a variation of RL that studies how a human can be integrated in the agent learning process. Human input can be either in the form of feedback or guidance. *Learning from Feedback* treats the human input as a reinforcement signal after the executed action [10]. *Learning from Guidance* allows human intervention to the selected action before execution, proposing (corrective) actions [2]. We argue that IRL and transfer learning techniques can be combined to enable efficient policy adaptation towards a new user. In the next sections, we present the proposed framework and the application of this to Robot Assisted Therapy.

3. PROPOSED FRAMEWORK

We propose an interactive learning and adaptation framework that enables an interactive agent to modify its policy in an online fashion. More specifically, the framework utilizes implicit feedback provided by the primary user to refine its learned policy towards the current user. However, ensuring a safe interaction while adapting the agent's behavior to a different user is a key challenge. Our framework supports the participation of a secondary user, as a supervisor, that can guide the interaction in its early steps, avoiding unsafe interactions. The goal of this framework is to enable agents to learn as long as they interact with primary and secondary users, adapting and refining their policy dynamically [23].

4. APPLICATION TO ROBOT ASSISTED THERAPY: YOGA TRAINING

In this section, we illustrate our proposed framework with a use case in Robot Assisted Therapy. More specifically, we present a yoga training system that dynamically adapts to different user abilities and needs. Yoga has been recently employed as a means for therapeutic rehabilitation, demonstrating its ability to promote physical strength, flexibility, respiratory and cardiovascular functions, as well as mental health, well-being, and overall quality of life [26]. For this paper, we plan to integrate physical routines from yoga with physiological feedback and robotic support as an interactive avatar.

In specific, we follow the following scenario. The user has to perform a set of five physical routines, prescribed by the therapist. Each routine consists of a predefined set of poses. The amount of time each pose must be performed can be modified to help the user complete the whole session successfully. A NAO robot demonstrates the prescribed exercises and the user is asked to perform the demonstrated exercise along with the robot. The system monitors the exercise execution and adjusts the session so as to help the user perform the exercises efficiently.

5. SYSTEM ARCHITECTURE

In this section, we present the architecture of our proposed system as shown in Fig.1. The robot demonstrates the prescribed exercises and asks the user to perform them. Based on its learned policy, the system adjusts the difficulty utilizing the motion data acquired by the Kinect sensor. Two additional communication channels (heart rate feedback and therapist guidance) are integrated to the adaptation module for the efficient policy refinement to the specific user.

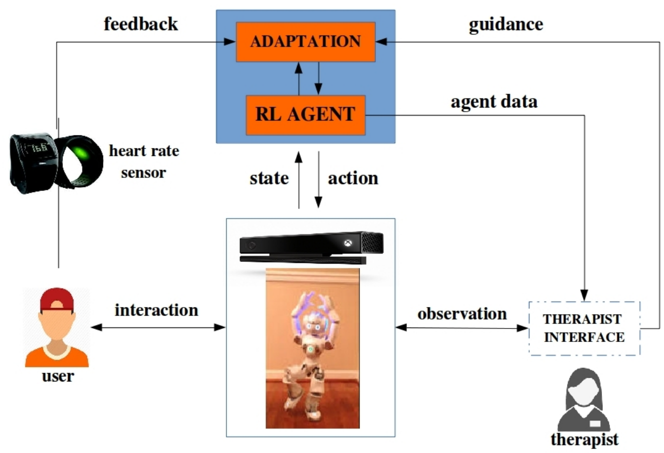

Figure 1: System Architecture.

5.1 The NAO platform

The NAO is a small humanoid robot developed by Aldebaran that can move with 25 degrees of freedom and is equipped with various sensors. It can be programmed visually using Aldebaran's Choreograph software, or by using C++ or Python with the provided SDK. The role of the NAO in this application is to demonstrate exercises to the user and ask them to perform them. The NAO robot has the ability to demonstrate basic yoga movements, depending on the complexity of the pose and how accurately it should resemble the human pose [12]. We investigate the exercises and poses that NAO can demonstrate with different type parameters (time, accuracy).

5.2 Hardware

Other hardware to be considered in this system, aside from the NAO, is foremost the Microsoft Kinect. The Kinect is a motion capture sensor with an RGB camera, infrared sensor and microphones that allow full body motion capture, facial and voice recognition. Heart rate is a relevant measurement that the Kinect v.2 is also capable of by detecting small fluctuations in the skin. More reliable, wearable heart rate monitors could also be used, such as a smart watch, which use infrared and LEDs that see flow of the blood in the wrist to measure heart rate. This would be slightly more cumbersome than solely a Kinect, but only slightly so, and with the benefit of increased accuracy and reliability, if needed.

5.3 The RL agent

The RL agent is responsible for the action selection of the system. An RL agent learns an optimal policy as it interacts with the user environment, receiving a reward for the transition of one state to another, by performing an action. The optimal policy is the mapping from states to actions that maximizes the expected total reward. In this section, we show the problem formulation as a Markov Decision Process (MDP).

The state space includes all the information needed by the agent to decide for its next action. The state space includes information about the session (which exercise is being demonstrated and the parameters of this exercise) and the user (user performance). The system considers the state input in order to decide for the next action. The agent must take the appropriate action, based on the current state, so as to help the user finish the session, maximize their performance and prevent them from quitting. The system can either adjust the time of the demonstrated movements, move to the next exercise, or encourage the user.

The goal of the RL agent is to find a policy that maximizes the expected return at the end of each episode. After each action, the agent receives a reward that evaluates the action selected based on the current state. The agent receives a reward for each transition, evaluating the current policy. The reward signal is a function of user performance as received by the human motion recognition and evaluation module output.

5.4 Human Motion Recognition and Evaluation

An important component of the proposed system is the human-activity recognition and evaluation module. That module will be responsible for providing performance related scores both to the RL module and the therapist.

Part of our ongoing work focuses on the development of a framework for human motion analysis by exploiting the advantages of the 3D point clouds offered by depth sensors (such as Microsoft Kinect). Recent literature has shown that depth maps have several advantages compared to traditional color images. For example, depth maps reflect pure geometry and shape cues, which can often be more discriminative than color and texture in many problems. Moreover, depth maps are insensitive to changes in lighting conditions and can secure privacy since color, texture and minor shape details are absent from the retrieve information[25].

Going towards that direction, we propose a two-step motion-analysis framework. Initially a step for real-time motion recognition will be applied. Deep Architectures [25, 4, 3] have dominated recent research in vision-based activity recognition showing state-of-the-art results in recognizing activities in semi-structured environments such as ours. As an additional step, we research on developing evaluation measures on the top of recognition. We discriminate evaluation measures in two basic categories. General measures and application-dependent measures. As general measures, we consider metrics related to the spatio-temporal information of the action. For example, time required to complete the action, range of motion when performing it and also time and range of motion deviation compared to the prescribed movement. Application-dependent measures can also be used to further tailor the system, as attention level [7] and the amount of extraneous movements.

5.5 Therapist Interface

In our setup, a therapist can be either present or observe the session remotely through a user interface. The user interface presents to the therapist information about the session and the user. This information can be used by the therapist to intervene to the interaction, when needed, Moreover, the therapist can guide the early interaction steps to set specific therapeutic goals. The interface also presents information about the agent policy, enabling the therapist to provide corrective actions, to either avoid an inappropriate action or to guide the agent's action selection.

5.6 Adaptation Module

The adaptation module is responsible for adapting a learned policy and refine it towards the current user. We propose the integration of IRL techniques, defining two extra communication channels; guidance and feedback. We argue that human knowledge and intentions can be communicated through these channels and can be utilized for an efficient and safe adaptation of the agent to the specific user.

Guidance is communicated to the agent through the therapist interface. It can either be a corrective action, or a set of proposed actions, based on the therapist therapeutic goals and expertise. This information is processed and integrated to the learning mechanism.

Feedback is implicitly provided by the user while exercising, through the heart rate. In research and clinical use, heart rate (HR) has consistently been used as a non-invasive and inexpensive method to calculate an individualâĂŹs cardiovascular responses to physical activity [11, 9]. As a true measurement in living persons, HR can effectively determine physical training work-loads through the recording of resting HR and an individual's max HR [19]. Additionally, HR can act as a universal metric that can adapted to individuals across age, gender, and habitual exercise status.

Implicit feedback can be considered as a personalization factor. The adaptation module uses this information to refine its current policy so as to adapt to the current user. On the other hand, guidance is provided by an expert user ensuring a safe interaction as the agent adapts to the user. Our ongoing work moves towards the combination of these two different channels and integration to the learning mechanism.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we presented an interactive learning and adaptation framework. The framework combines interactive learning techniques and utilizes them to refine a learned policy towards a new user. We argue that implicit feedback, provided by the primary user and guidance from a secondary expert user can be integrated to the learning mechanism, aiming at a tailored and safe interaction. As an ongoing work, the next step is to study each module separately and then integrate them. Considering the adaptation framework, we need to investigate how feedback can be used to efficiently modify a learned policy. Feedback must be handled as a policy modifier and not as an additional reinforcement signal, since it may not alter the policy. Guidance can be seen as a human-guided exploration mechanism; [18] However, therapist interventions should also provide additional information to the agent, facilitating the adaptation. A significant aspect of guidance is the workload of the therapist. Therapist's interventions should reduce as the agent learns, indicating that the agent converges to an optimal policy. Active Learning methods can be used to learn, based on state information (i.e., state uncertainty and importance), when the therapist should intervene.

7. ACKNOWLEDGMENTS

This material is based upon work supported by NSF under award numbers CNS 1338118, 1035913 and by the educational program of NCSR and UTA. Also this work is co-financed by the European Union (European Social Fund ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program:Thalis- DISFER

8. REFERENCES

[1] J. Broekens. Emotion and reinforcement: affective facial expressions facilitate robot learning. In *Artifical Intelligence for Human Computing*. Springer, 2007.

[2] F. Cruz et al. Interactive reinforcement learning through speech guidance in a domestic scenario. In *Proc. IJCNN*. IEEE, 2015.

[3] T. Dobhal et al. Human activity recognition using binary motion image and deep learning. *Procedia Computer Science*, 2015.

[4] J. Donahue et al. Long-term recurrent convolutional networks for visual recognition and description. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[5] N. Giullian, D. Ricks, A. Atherton, M. Colton, M. Goodrich, and B. Brinton. Detailed requirements for robots in autism therapy. In *Proc SMC 2010*. IEEE.

[6] M. A. Goodrich et al. Incorporating a robot into an autism therapy team. *IEEE Intelligent Systems*, 2012.

[7] J. Hernandez et al. Measuring the engagement level of tv viewers. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*. IEEE, 2013.

[8] S. Janarthanam and O. Lemon. Adaptive generation in dialogue systems using dynamic user modeling. *Computational Linguistics*, 2014.

[9] J. Karvonen and T. Vuorimaa. Heart rate and exercise intensity during sports activities. *Sports Medicine*, 1988.

[10] W. B. Knox and P. Stone. Reinforcement learning from simultaneous human and mdp reward. In *Proc. AAMAS*, 2012.

[11] A. Kolus, D. Imbeau, P.-A. Dubé, and D. Dubeau. Classifying work rate from heart rate measurements using an adaptive neuro-fuzzy inference system. *Applied Ergonomics*, 2016.

[12] D. Lee, W.-h. Yun, C. Park, H. Yoon, and J. Kim. Analysis of children's posture for the bodily kinesthetic test. In *Ubiquitous Robots and Ambient Intelligence (URAI), 2015 12th International Conference on*. IEEE.

[13] C. Li, Z. Rusák, I. Horváth, L. Ji, and Y. Hou. Current status of robotic stroke rehabilitation and opportunities for a cyber-physically assisted upper limb stroke rehabilitation. In *Proceedings of the 10th international symposium on tools and methods of competitive engineering TMCE 2014*. TMCE.

[14] R. Loureiro, C. Collin, and W. Harwin. Robot aided therapy: challenges ahead for upper limb stroke rehabilitation. In *Proceed of Intl Conf on Disability, Virtual Reality and Assoc Tech*, 2004.

[15] H. Modares, I. Ranatunga, F. L. Lewis, and D. O. Popa. Optimized assistive human–robot interaction using reinforcement learning. 2015.

[16] R. Picard. *Affective Computing*. MIT Press, 1997.

[17] A. Ramachandran and B. Scassellati. Developing adaptive social robot tutors for children. In *2015 AAAI Fall Symposium Series*, 2015.

[18] E. Senft, P. Baxter, and T. Belpaeme. Human-guided learning of social action selection for robot-assisted therapy. In *4th Workshop on Machine Learning for Interactive Systems*, 2015.

[19] H. Tanaka, K. D. Monahan, and D. R. Seals. Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology*, 2001.

[20] A. Tapus, C. Tapus, and M. Matarić. Long term learning and online robot behavior adaptation for individuals with physical and cognitive impairments. In *Field and Service Robotics*. Springer, 2010.

[21] M. E. Taylor and P. Stone. Transfer learning for reinforcement learning domains: A survey. *The Journal of Machine Learning Research*, 2009.

[22] A. L. Thomaz and C. Breazeal. Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance. In *AAAI*, 2006.

[23] K. Tsiakas. Facilitating safe adaptation of interactive agents using interactive reinforcement learning. In *Companion Publication of the 21st International Conference on Intelligent User Interfaces*. ACM.

[24] K. Tsiakas, M. Huber, and F. Makedon. A multimodal adaptive session manager for physical rehabilitation exercising. In *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, 2015.

[25] P. Wang et al. Deep convolutional neural networks for action recognition using depth map sequences. *arXiv preprint arXiv:1501.04686*, 2015.

[26] C. Woodyard et al. Exploring the therapeutic effects of yoga and its ability to increase quality of life. *International journal of yoga*, 2011.

[27] W. Zemzem and M. Tagina. A novel exploration/exploitation policy accelerating learning in both stationary and non-stationary environment navigation tasks. *International Journal of Computer and Electrical Engineering*, 2015.